Please amend the claims as follows. This listing of claims will replace all prior versions, and

Listings of Claims in the application:

<u>Listing of Claims</u>:

1.-8. **(cancel)**

9. (Currently Amended) A computer implemented method comprising:

monitoring usage of a computing resource utilized by a workload wherein said monitoring is performed by a process within a user space and the process monitors the user space only, the workload includes a plurality of running processes, the plurality of running processes are a subset of all processes that are running in the user space;

responsive to exceeding a limit on utilization of said computing resource, decreasing usage of said computing resource by said workload, <u>said computing resource comprises physical memory and said decreasing usage of said computing resource comprises paging a portion of said physical memory assigned to said workload out of said physical memory and said decreasing usage does not halt operation of said workload.</u>

10.-11. **(Cancel)**

12. (Currently amended) The method of claim <u>9</u> [[11]] wherein said portion of said physical memory comprises a least recently used portion of said physical memory assigned to said workload.

13. **(Cancel)**

14. (Original) The method of claim 9 wherein said decreasing usage is initiated by a process of said workload.

15. (Original) The method of claim 9 wherein said process that performs said monitoring is not an operating system kernel process.

16. (Currently Amended) A computer implemented method for memory management of a workload from within the workload comprising:

accessing a list of memory pages assigned to said workload in a user space, the workload includes a plurality of running processes, the plurality of running processes are a subset of all processes that are running in the user space;

responsive to a request from a first process of said workload for memory which exceeds a predetermined memory limit for said workload, selecting a plurality of memory pages from said list of memory pages, wherein the plurality of memory pages includes least recently used memory pages assigned to the workload; and

initiating a second process within the user space to page out said plurality of memory pages wherein at least a portion of said workload continues to operate subsequent to said initiating.

17. (Original) The method of claim 16 wherein said accessing, selecting and initiating are performed by said second process within said workload.

18. (Original) The method of claim 16 wherein said second process is not an operating system kernel process.

19. (Original) The method of claim 16 wherein said second process is loaded into a user space.

20. (Original) The method of claim 16 wherein said plurality of memory pages comprises memory pages that are least recently used.

21. (Original) The method of claim 20 wherein said plurality of memory pages comprises memory pages that are least recently used by said workload.

22. (Original) The method of claim 20 wherein said page out of said plurality of least recently used memory pages reduces a number of memory pages assigned to said workload to below said memory limit.

23. (Original) The method of claim 20 wherein said plurality of least recently used memory pages comprises the minimum number of memory pages to reduce said number of memory pages assigned to said workload below said memory limit.

24. (Cancel)

25. (Original) The method of claim 16 wherein said initiating is not performed by an operating system kernel process.

26.-35. (Cancel)

36. (Currently amended) A computer implemented method comprising:

accessing memory usage for a workload and examining page usage for each process of said workload, the workload exists within a user space and includes a plurality of running processes, the plurality of running processes are a subset of all processes that are running in the user space;

aggregating usage of said each process to determine an aggregate usage for said workload;

if said aggregate usage does not exceed a memory utilization limit for said workload, repeating said accessing and aggregating for a next workload;

if said aggregate usage exceeds said memory utilization limit for said workload, determining least recently used pages based on accessed bits associated with said workload;

if said aggregate usage exceeds said memory utilization limit for said workload, supplying a range of least recently used pages in a system call to an operating system kernel for evicting said range of least recently used pages to reduce resource usage by said workload; and

retaining at least partial operation of said workload during said page evicting <u>such that</u> <u>the at least partial operation does not stop said workload</u>.

37. (Original) The method of claim 36 wherein said determining and said supplying occur in a plurality of user space processes.

38. (Currently Amended) A computer implemented method of managing computer resources over a plurality of workloads, said method comprising:

for each workload of said plurality of workloads, monitoring respective workload resource usage against a respective allotment of each workload;

determining a range of computer resources to page out for each workload whose resource usage exceeds its respective allotment; and

initiating a paging out operation of said range of computer resources and wherein said monitoring, said determining and said initiating all occur within a process of user space,

wherein each of the plurality of workloads exists within a user space and includes a plurality of running processes, the plurality of running processes are a subset of all processes that are running in the user space<u>, and paging out said range of computer resources and</u> <u>wherein each workload whose resource usage exceeds its respective allotment remains</u> <u>partially operable during said paging out of its respective range of computer resources</u>.

39. (Original) The method as described in claim 38 wherein said determining comprises determining least recently used pages for each workload whose resource usage exceeds its respective allotment.

40. (Original) The method as described in claim 38 wherein said process is situated within a workload of said plurality of workloads.

41.-53. (Cancel)